

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2001, 19:40:41 ; Search time 181.41 Seconds
(without alignments)
18724.010 Million cell updates/sec

File: US-08-153-397a-1
Perfect score: 3962
Sequence: 1 CGGGCTGAGACTGGGCTGA.....AAAAAAAAAACCGGAATTC 3962

Scoring table: IDENTITY_NUC 1
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3962	100.0	3962	AA93785	Human mammary carc
2	3960.4	100.0	3962	AA92520	Human mammary carc
3	3960.4	100.0	3962	AA92522	Human mammary carc
4	3617.6	91.3	3754	AA084782	Protein-tyrosine-k
5	2783.6	70.3	2861	AA090501	Human pancreatic c
6	2783.6	70.3	2861	AA090501	Human pancreatic c
7	642	16.2	3096	AA048292	Human colon cancer
8	642	16.2	3157	AA092521	Discolidin domain r
9	642	16.2	3157	AA092523	Human colonic aden
10	642	16.2	3157	AA092523	Human colonic aden
11	639.8	16.1	3120	AAV55895	CKK-2, a human mam
					Receptor protein t

12	639.8	16.1	3120	AAV55317	Receptor protein t
13	544.2	13.7	2725	AAH95505	Human protein enco
14	453.4	11.4	2128	AA064158	Partial coding seq
15	324.4	8.2	408	AA000624	Human secreted pro
16	236	6.0	272	AA022477	Human gene signatu
17	228.4	5.8	463	AA11731	Probe #1664 for ge
18	228.4	5.8	463	AA133038	Probe #1724 used t
19	228.4	5.8	463	AA101659	Probe #1650 used t
20	182.2	4.6	2820	AAV51456	gd.trkA fusion use
21	180.6	4.6	2301	AAV20445	Human c-trk oncoge
22	180.6	4.6	2301	AAV09303	Human trk oncogene
23	180.6	4.6	3060	AAV51457	gd.trkb fusion use
24	180.6	4.6	3194	AAV00689	Human trkb recepto
25	180.6	4.6	3194	AAV00689	Human trkb recepto
26	180.6	4.6	3707	AA099277	Human neurotrophic
27	180.2	4.5	2526	AA069029	Porcine trkC gene.
28	178.6	4.5	2526	AA028668	Encodes adult porc
29	178.6	4.5	2526	AA034581	trkC clone in PFLI
30	178.6	4.5	2940	AAV51458	gd.trkC fusion use
31	177	4.5	2750	AAV87599	DNA coding for neu
32	158	4.0	4092	AAV70230	Probe #10877 for g
33	151	3.8	175	AA120944	Probe #10877 used t
34	151	3.8	175	AA16188	Partial sequence o
35	151	3.8	175	AA106655	Probe #6646 used t
36	140	3.5	2376	AAQ34582	Mouse muscle-local
37	139	3.5	2208	AAV16347	Encodes mouse trkC
38	139	3.5	2604	AAV16349	Degenerate rat ins
39	138.4	3.5	2376	AA028669	Nsk2 receptor gene
40	138.4	3.5	2376	AA069030	Mouse receptor tyr
41	138	3.5	3900	AA068343	Human insulin rece
42	137.4	3.5	3257	AAV14352	
43	137.4	3.5	3257	AAV8546	
44	135.2	3.4	1212	AAV59295	
45	135.2	3.4	4723	AAH50568	

ALIGNMENTS

RESULT 1	
AA93785	
ID	AA93785 standard; CDNA; 3962 BP.
AC	AA93785;
DT	16-FEB-1998 (first entry)
XX	
DE	Human mammary carcinoma kinase 10 (MCK-10) cDNA sequence.
XX	
KW	Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW	proliferative disease; cancer; insulin receptor family;
KW	tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW	neurological disorder; aberrant expression; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
XX	
XX	Location/Qualifiers
XX	321..3077
XX	/*tag= a
XX	
XX	US5677144-A.
XX	
XX	14-OCT-1997.
XX	
XX	08-NOV-1994; 94US-0336343.
XX	
XX	16-NOV-1993; 93US-0153397.
XX	
XX	(ALVE/) ALVES F H E.
XX	(ULR/) ULLRICH A.
XX	
XX	Alves FHE, Ullrich A;
XX	

QY 1741 ACCGCCAGTCTAGAGAGCACCCTTACAGAGCCCGGCTGTGGGAATCCG 1800
|||||
Db 1741 aacgcacagctctctagagagcaacccctacagagagcccgctctggaatccgc 1800
|||||
QY 1801 CCCACTCCGCTCCCTGTGTCCCAATGGCTCTGCTGTGCTCTCAATCCAGCTTAC 1860
|||||
Db 1801 cccactccgctccctgtctcccaatgctcgtctgtctgtctcccaatccagctaac 1860
|||||
QY 1861 GCTCTCTTGTGGCCACTTACGCCCTCCCTCGAGGCCGGGCCCCCCACACCCGCT 1920
|||||
Db 1861 gccctctctgtgccaactacgcccgtccctctgagcccgccccccacacccgct 1920
|||||
QY 1921 GGGCCAAACCCACCAACCCAGGCTACAGTGGGAGTATATGAGAGCTTGAAAGCCAG 1980
|||||
Db 1921 gggccaaacccacacacacacagcgtctacagtggaactataatgagcctggaagccag 1980
|||||
QY 1981 GCGGCCGCTTCTGCCCCCACCCTCCCGAGAACAGCGTCCCATATGCGGAGGCTGACA 2040
|||||
Db 1981 gcgcccgctctctgccccacactccacagacagctcccatatctcgaagctgaca 2040
|||||
QY 2041 TTGTACCTCTCAGAGGCGTCAACGGGGGCAACACTATGCTGTGCTGCTCACTGCCCCAG 2100
|||||
Db 2041 ttgtacctctcagagcgctcaccggggcaacactatgctgtgctgactgccccag 2100
|||||
QY 2101 GGGCAGTCCGGGAGTGGGCCCCCAGAGTGGATTTTCCCTCGATCTCGACTCCGCTTCAAG 2160
|||||
Db 2101 gggcagtcgggagtgggccccccagagtgagatctccctcgactctgactccgctcaagg 2160
|||||
QY 2161 AGAAGCTTGGGAGGAGGCGCAGTTTGGGAGTGCACCTGTGAGAGTGCACAGCCCTCAG 2220
|||||
Db 2161 agaaagctctgagagggcagagcttgaggagtgacacgtgtgagagtgacagccctcag 2220
|||||
QY 2221 ATCTGTACAGTCTTGAATTTCCCTTAATGTCGTAAAGGACACCTTGTGTGAGTGTG 2280
|||||
Db 2221 atctgtcagctctgtatctcccttaattgtgcttaaggagacacacttgctgtagtgtg 2280
|||||
QY 2281 TCAAGATCTTACGGCCAGATGTCACCAAGAAATGCCAGCTTCTCTTGTTCACAGGAATG 2340
|||||
Db 2281 tcaagatcttaccgagccagatgcccacagaaatgccaagctctctctgtctccaggaatg 2340
|||||
QY 2341 ATTTCGTAAGAAGTGAAGTATGTCAGGCGTCAAGAGCCCAACATCAATTTGGGCTGC 2400
|||||
Db 2341 atttcctgaagaaggggtgaagatactctgagggcccaagagcccaaatctctgctgc 2400
|||||
QY 2401 TGGCGCTGTGTGACGAGACGCCCTCTGTGATGATTACTGATGATGAGAGAGCGCG 2460
|||||
Db 2401 tggcgctgtgtgacgagacgccctctgtcatgattactgactatactgagaaagcgcg 2460
|||||
QY 2461 ACCTCAACAGTTCCTCAGTGCACCAAGCTGAGAGCAAGAGCAGCGAGGGGCCCTG 2520
|||||
Db 2461 acctcaacagttctctcagtgccacacagctgagagcaagcgagcgaggggccctg 2520
|||||
QY 2521 GGGACGGGCGAGGTGGCGACGAGGGGCCACATACGCTACCAATGCTCTGTGATGTGGCAG 2580
|||||
Db 2521 gggacgggcgaggtgctgacaggggcccacatcagctacccaatgctgtgaaatggaag 2580
|||||
QY 2581 CCCAGATCCGCTCCGCGATCGCTATCTGGCCACTCAATTGTGATCGGAGACTGG 2640
|||||
Db 2581 cccagatccgctccgctcgcatctgtatctgccaacactcaactttgttcaatcggaacctg 2640
|||||
QY 2641 CCACGGGGAAGTCTCAGTTGGGAAAATTTCAACATCAAAATCGACAGACTTTGGCATGA 2700
|||||
Db 2641 ccacggggaactctcagttgaggaaaatttcaacatcaaaatcgacagactttgcatga 2700
|||||
QY 2701 GCCGGAACCTTATGCTGGGAGTATTAACGATGAGGCGCGGCGAGTCTGCGCCATCC 2760
|||||
Db 2701 gccggaaccttatactgtggaactataacagtgtagagggcggaagctgtcgccatcc 2760
|||||
QY 2761 GCTTGATGGCTGTGGAGTGCATCTCATGGGAAAGTTCAAGACTGCGAGTGTGGG 2820
|||||
Db 2761 gcttgatggcctgtggagtgcatctcatctgtaggaagttaacagactgtgagtgagtgagg 2820
|||||

QY 2821 CCTTGTGTGACCTCTGTGGAGAGTCTGATGCTGTGAGGGGCCACCCCTTTGGGACG 2880
|||||
Db 2821 ccttgtgtgacacctctgtgtggagagtgctgatgtctgttagggccagcccttggcgagc 2880
|||||
QY 2881 TCACCGAGAGAGGATCATCGAAGACGGGGAGTTCCTCCGGGACACAGGCGCGCAG 2940
|||||
Db 2881 tcaccgagagagagtcatactgaaagcgagggttctccgggacagggccggcgag 2940
|||||
QY 2941 TGTACTGTCCCGGCGGCTGCTGCGCCAGAGGCTATATGAGCTGATGCTGCTGCT 3000
|||||
Db 2941 tgtactgtcccgcgccgctctgctgcgagggcctataatgactgactgtcttggtct 3000
|||||
QY 3001 GGAAGCGGAGTCTGAGCAGCAGCACCTTTTCCAGCTGATCGGTTCTCTGCGAGAG 3060
|||||
Db 3001 ggaagcggagctgtgacagccacccctttccagagtgatctgctccggagag 3060
|||||
QY 3061 ATGCACTCAACAGGATGTAATCACATCCAGCTGCCCCCTCCCTCAGGAGTATCCAG 3120
|||||
Db 3061 atgcaactcaacaggtgtgaatcacacatccagctgccccctccacagggagtatcag 3120
|||||
QY 3121 GGGAAAGCCAGTGAACACTAAACAGAGACAAAGGACGCTGCGCTTCCCTCCCGGA 3180
|||||
Db 3121 gggaaagccagtgaacactaaacagagacaaatgacacatgacccctctccctccga 3180
|||||
QY 3181 CAGCCATCACCTTAATAGAGCAGTGAAGTGAAGTGGGCTGGGCCAACCCAGGAG 3240
|||||
Db 3181 cagccatcacacttaataagagcagtgagactgagtgagtggtgtggccacacagggag 3240
|||||
QY 3241 CTGATGCCCTTCTCTCCCTCTCTGACACACTCTCATGTGCCCTTCTGTCTCTTCC 3300
|||||
Db 3241 ctgatgcccttctctccctctctctgacacactctcatgtgcccttctgtctcttcc 3300
|||||
QY 3301 TAGAAGCCCTGTGCGCCACCCAGCTGCTCTGTGAGTGGATTCCTTCACACCTCTCT 3360
|||||
Db 3301 tagaagccctgtgctgccccacacagctgctctgtgagtgagctctctccacccctct 3360
|||||
QY 3361 AGCCATCCCTTGGGAGAGTGGGGAATATAGATAGACACTGAGCATGGCCCATTTG 3420
|||||
Db 3361 agccatcccttgggagaggggtgggaataatagagataagacactgagacatgagcccat 3420
|||||
QY 3421 GAGCACTGGGCGCCACTGACACACACTGATTCCTGAGAGGTGGCTGGCCCAAGCTTC 3480
|||||
Db 3421 gagcactgggcccccaactgagacacactgattctctgagaggtgtgctgcgccccagctc 3480
|||||
QY 3481 TCTCTCCCTGTCAACAACAATGAGACCCCACTGGCTGAGAACTCTGGGGGTAGAGAGCAAGA 3540
|||||
Db 3481 tctctccctgtcaacaacaatgagaccccaactggctgagaaactctgggggttagagagaa 3540
|||||
QY 3541 AGGAGAGGAAATGTTCTCTGTGCTGCTCTCTGTAACCTTGTCTCAAGCTTGGGCTTCTC 3600
|||||
Db 3541 aggagaggaatgtttctctgtgtcctgtcctctgtactgttcccaagcttgggtcttc 3600
|||||
QY 3601 CTCCTCATCACTGAAACACTGGAAGCTGGGAGTGGGCCGCCCAAGCCCTCACTACCC 3660
|||||
Db 3601 ctctctcatcactgaaacactggaagctggggtagcccgccccaagccctcagctaaccc 3660
|||||
QY 3661 CCACCTCCCACTTGAGCTGTAGCTGATGAACTCTCTCAACCTCATAGCTTGTGTGGAG 3720
|||||
Db 3661 ccactctccacttgagctgttagctagaactctctcaagccatcagcttctctgtgag 3720
|||||
QY 3721 TAAATATTGGGATTTGGGGAAGAGGAGGAGGAGGCGCCATAGGCTTGGGGTGGAGATC 3780
|||||
Db 3721 taatattggatattggggggaagaggaaggaagcccatagccttgggggttggagatc 3780
|||||
QY 3781 TCTAGTGTAGTGGCACATTGATTTTCTATATCACTTGGGGTTGTATATTTTGGGG 3840
|||||
Db 3781 tctagtgtagctgacacatgatttctataacacttgggggttcttactatcttttggg 3840
|||||
QY 3841 GGAAGAGACAGATTTTATACATTAATATGAGACCTAGCTTGAAGCAATTTTATCCCT 3900
|||||
Db 3841 ggaagagacagatcttctacactaataatagacctagcttgaagcaatlttaatccct 3900
|||||
QY 3901 GCACTAGGACGATTAATAAAGTTGAGTTTCCACAAAAAATAAAAAACCGGAAT 3960
|||||


```

FT   CDS                               321..3080
FT   /*tag= a
FT   misc_difference 2315
FT   /*tag= b
FT   /note= "some clones have AA deletion here"
XX   PN   W09514089-A.
XX   PD   26-MAY-1995.
XX   PE   16-NOV-1994; 94WO-EP03799.
XX   PR   16-NOV-1993; 93US-0153397.
XX   PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX   PI   Alves FHE, Ullrich A;
XX   WP1; 1995-224055/29.
XX   P-PSDB; AAR75504.
XX   PT   New nucleic acid encoding CCK-2 receptor tyrosine kinase - and
XX   derived vectors, transformed cells, proteins and antibodies, useful
XX   for diagnosis and treatment of proliferative and nervous system
XX   diseases and for screening modulators
XX   PS   Disclosure; Page 67-69; 115pp; English.
XX   CC   cDNA prep'd. from human breast cancer cell line MCF7 (ATCC HTB22) was
XX   used in a PCR with two degenerate oligo primer pools based on
XX   conserved sequences of the kinase domain of receptor tyrosine
XX   kinases. One clone, designated MCK-10, was identified as novel RTK.
XX   The PCR fragment was used to screen a lambda gtl1 library of human
XX   fetal brain cDNA. Several overlapping clones were identified. The
XX   composite of these cDNA clones is given in A093252 and the deduced AA
XX   sequence in AAR75504. Some of the clones had a deletion of 6AA at posn.
XX   2315 in the MCK-10 sequence. MCK-10 has all the characteristics of
XX   a receptor PRK (see AAR75504 FT). Screening of human placental library
XX   yielded two cDNA clones. One of the clones isolated from the human
XX   fetal brain library contained an additional 18 nts in the TK
XX   domain. The MCK-10 splice isoforms have been designated MCK-10-1
XX   (with an additional 111 bp between nts 1832 and 1943); MCK-10-2
XX   (without any insertions); MCK-10-3 (with the additional 111 bps and
XX   18 bp in the TK domain); and MCK-10-4 (with the additional 18 bp).
XX   The predicted mol. wts. of MCK-10-1 and MCK-10-2 proteceptors are
XX   104.13 and 97.17 kD respectively, and can thus be subdivided into a
XX   34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that
XX   contain the TK homology and alternative splice sites.
XX   SO   Sequence 3962 BP; 735 A; 1235 C; 1181 G; 811 T; 0 other;

Query Match      100.0%; Score 3960.4; DB 16; Length 3962;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3961; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0Y   1 CGGGCCCTGAGACTGGGGTGACTGGGACCTTAAGACAAATCTGAGCTGAGAGCCCCCGACAG 60
Db   1 CGGGCCCTGAGACTGGGGTGACTGGGACCTTAAGACAAATCTGAGCTGAGAGCCCCCGACAG 60
0Y   1 CTGCTCTGGGAGAGCGGCTCCCGACACCCGAGCCCGCGGGCGCTCCCGTCCCGGCTC 120
Db   1 CTGCTCTGGGAGAGCGGCTCCCGACACCCGAGCCCGCGGGCGCTCCCGTCCCGGCTC 120
0Y   121 CGGGCTCTGAGCTCCCTCCGCGCTCCCGCGCCCTCGCCCGCCGCGCCGCAAGAGCCCGCT 180
Db   121 CGGGCTCTGAGCTCCCTCCGCGCTCCCGCGCCCTCGCCCGCCGCGCCGCAAGAGCCCGCT 180
0Y   121 CCGGCTCTGCTCTCCCTCCGCTCCCGCGCCCGCCGCGCCGCGCGCGCGCGCGCGCGCT 180
Db   121 CCGGCTCTGCTCTCCCTCCGCTCCCGCGCCCGCCGCGCGCGCGCGCGCGCGCGCGCT 180
0Y   181 CCGCGGCTGGAAGCGCTGGGCTGCGCGGGAAGAGCATGAGAGTGCTGTGAAGTGCGCTAT 240
Db   181 CCGCGGCTGGAAGCGCTGGGCTGCGCGGGAAGAGCATGAGAGTGCTGTGAAGTGCGCTAT 240
0Y   241 TCATGAGCGGATGGGGTTGGACTGGAAGGAATGCCAGAGAGTGTGCGCCCACTCCCTTA 300
Db   241 TCATGAGCGGATGGGGTTGGACTGGAAGGAATGCCAGAGAGTGTGCGCCCACTCCCTTA 300

```

```
Db 241 tcaactgacgatagggttggaacttgaaagaaatgccaagaatgctgcaccccccctta 300
QY 301 GGGCCGAGGAGTCAAGGAGCTATGGAGACGAGGCCCTGTCACTTATCTTACTGCTGCTCTT 360
Db 301 gggccgagggatcagagactatggaacagagcccctgtcaacttactgcgcgcctct 360
QY 361 TGGTGGCAAGTGGAGATCTACATGAGAGGACATTTGATCTGCTGCAAGTGGCCGTATG 420
Db 361 tgggagcaagggatgctgactgacatgaaaggacatttgatccctgcagtgccgtatg 420
QY 421 CCCGTGGGATCAGAGACCGGACCATCCAGACAGTACATCTCTCTTCCAGCTCTGTGT 480
Db 421 cccctgggatgcaagaccgagacacatccagacagtgacatctctctccagctcctgt 480
QY 481 CAGATTCCACTGCGCCCGCCGACACbbaagttggagagacagatgacggagtgggcctgt 540
Db 481 cagattccaactgcgcgcgcgcagcaggttgagagacagtgagcgaggatgggctgtgt 540
QY 541 GCGCCGAGGGTCCGATGTTTCCCAAGGAGGAGGATCTTGAGAGGTGGATATCAACGAC 600
Db 541 gcccgcagggctcggtgttcccaaggagagagatctgcaagtggacltcaacagac 600
QY 601 TCCACTGTGTGGCTCTGGTGGGACCCAGGAGCGCATGCGGGGCTGTGGGCAAGAGT 660
Db 601 tccacctgtgtctgtgtggcaaccagggacgcatgcccggggcctgtggcaaggagt 660
QY 661 TCTCCCGAGGTAACCGGCTGCTTACTCCGGGATGTCGGCGCTGATGGGCTGGAAG 720
Db 661 tctcccggaactacgcgctgctgtactcccggaatgtgcgcgcgtgagtggcgtaagg 720
QY 721 ACCGCTGGGTCAGAGAGTGTATCTCAGGCAATGAGGACCCGAGGAGATGGTGCTGAAG 780
Db 721 accgctgggtcagagagtgatctcccaaggacaaatgagaccgagagagatgtgcgaag 780
QY 781 ACCTTGGGCCCCCATAGTGTGCCCACTGTGCTCCCTTCAACCCGGGCTGACCGGATCA 840
Db 781 accttgggcccccatagtgtgcgcgactgtgtcttaccgcccggtctgacgggtca 840
QY 841 TGAATGCTGTCTGCGGGGATGAGCTCTATGAGCTCCCTGTGAGGAGATGAGTCTGTCT 900
Db 841 tgaatgctgtctgcggggatgagctctatgagctccctgtgagagatgactccgtctt 900
QY 901 ACACCGGCGCTGAGGGGACAGCAATGATTTATATGAGGCGGCTGTACTCAACACTCCA 960
Db 901 acacggcgccctgagggacaaatgatatcttgagccgtgttacctcaacgaactcca 960
QY 961 CCTATGACGAGACATACCGTGGGCGAGCTGCAGTATGGGGCTGTGGGACGTCAGATG 1020
Db 961 cctatgacgagacataccgtgtgggactgcaatgctgagttcgggccagactggcagatg 1020
QY 1021 GTGTGGTGGGCTGATGACTTTAGAAAGTACGAGCTGCGGGCTGTGGCAGGCTATG 1080
Db 1021 gtgtgggtgggctgatgactttagaaagtacgagctgcgggctgtggcaggctatg 1080
QY 1081 ACTATGTGGGATGAGACAAACACAGCTTTCACAGTGCCTATGTGAGATGAGATTGAGT 1140
Db 1081 actatgtgggattgagacaaacacagcttctccagtgtgctatgtgagatgagattgagt 1140
QY 1141 TTGACGGGCTGAGGGGCTTCCAGGCTATGAGGTCACGTAAACAATGACACGCGTGG 1200
Db 1141 ttgacgggctgaggggcttccaggctatgaggtccacgttlaaacaatgacacgctgg 1200
QY 1201 GAGCCGCTGTCCCTGCGGGGTGAATGTGCTTCGCGCGTGGCCCTGACATGGCCTGGG 1260
Db 1201 gagccgctgtccctgcggcggggtgaaatgtgcttcgctcgcgctggccctgcatggcggg 1260
QY 1261 AGGGGAGGCCATGCGGCACACACTAGGGGGCAACTGGGGGACCCAGAGCCCGGCGTG 1320
Db 1261 aggggagggccatgcggcacacactaggggcaacctgggggagcccccagggccggctg 1320
QY 1321 TCTCAATGCCCTTGGCGGCGGTGTGCTGTGCTTGTGAGTGGCGGCTCTTGTTCGGG 1380
Db 1321 tctcaatggccttggcgcggtgtgctgtgtgcttctgagtagcgcgcttctcttggcg 1380
Db 1381 GGGCCTGTGTACTCTTCAAGCAAAATCTCCTTATCTGTGATGTGGTGAACAAATCTCTTC 1440
Db 1381 gggcctgtgtactcttcaagcaaaatctccttcatctctgagtgtggaacaattctctcc 1440
QY 1441 CGGCACTGGGAGGACCTTCCCGCCAGGCCCTTGATGGCGGCTGTGGCCACTCCACCA 1500
Db 1441 cggcaactgggagggaccttcccgccagggcccttgatggcggctgtggccactccacca 1500
QY 1501 ACTTACGACGCTTGGAGCTGAGAGCCAGAGGACAGGCGCGGAGCAAGGGCGAGGGGA 1560
Db 1501 acttacgacgcttggagctgagagccagagagccagagccgtgccaaggccgaaggga 1560
QY 1561 GCCCGACCGCATCTCATCGGCTGCTGATGAGCCATCATCTGCTCTCATATCA 1620
Db 1561 gcccgacccgatctcatcggctgctgagtggagccatctcctgctcctgcgtcatca 1620
QY 1621 TTGGCCTCATGCTCTGCGCGCTGACATGGCGCAGGCTCCTCAGCAAGGCTGAACGAGG 1680
Db 1621 ttggcctcatgctctgcgcgctgacatggcgcgaggtcctcagcaaggctgaacgagg 1680
QY 1681 TGTGGAAGAGAGGTGAGAGGTTCACTCTCTGCTGGGGAACAATCATCATCAACA 1740
Db 1681 tgtggaaagagaggtgagaggttcaactctctgctcctggggaacaatactcatcaaca 1740
QY 1741 ACCGGCCAGGTCCTTAGAGAGACCCGCCCTACAGAGAGCCCGGCGCTGTGGAAATCCG 1800
Db 1741 accggccaggtccttagagagacccgccctacagagagcccggtcgtggaaatccg 1800
QY 1801 CCCACTCGGCTCCCTGTGTGCCCAATGGCTCTGCGTGTGCTTCATTCAGGCTTAC 1860
Db 1801 cccactcggctccctgtgtgcccaatggctctgcggtgtgcttctcattcagactacc 1860
QY 1861 GCCCTCTTGCGCACTTAAGCGCGCTCCCGCTCGAGGCGCCGGGCCCCCAGACCCGCT 1920
Db 1861 gccctcttgcgcaacttaagcgcgctcccgctcgagggccgggcccccaacacggct 1920
QY 1921 GGGCCAAACCCACAAACCCAGGCTACAGTGGGAGACTATATGAGAGCTGTGAAGCCAG 1980
Db 1921 gggccaaacccacaaacccaggctacagtgggagactatattgagagctgtgaagccag 1980
QY 1981 GGGCCCGCTCTGCGCCCACTCCCAAGACAGCGTCCGCCATTATGCCAGGCTGACA 2040
Db 1981 gggcccgctctgcgcccaactcccaagacaggtgcccaattatgcccaggtgaca 2040
QY 2041 TTGTATACCTCGAGGGGCTCACCGGGGACACACTATGCTGTGCTCAGTCCGCCAG 2100
Db 2041 ttgtatacctcgaggggctcacggggacacactatgctgtgctcagtgctgacgtcccccag 2100
QY 2101 GGGCAGTGGGAGTGGGCGCCCGCAGAGTGGATTTCCCTCGATCTGAGTCCGCTCAAG 2160
Db 2101 gggcagtgggagtgggcgcccgccagagtgatcttcccctcgatctcgactccgctcaag 2160
QY 2161 AGAAGCTTGGGAGGCGCAGTTTGGGAGTGCACCTGTGAGAGTCAGACGCCCTCAAG 2220
Db 2161 agaagcttgggagggcgagttgggagtgagacctgtggaaggtcgacaagccctcaac 2220
QY 2221 ATCTGGTATGCTTGTATTTCCCTTAATGTGCGTAAAGGACACCTTGTGCTGTAGTGTG 2280
Db 2221 atctggatgcttgtatttcccttaatgtgctaaaggacaccttgtgtgtagtctg 2280
QY 2281 TCAAGATCTTACGGGACAGATGCCAAGAATGGCAGATTCTCTGTGTCAGAGATG 2340
Db 2281 tcaagatcttaccgggacagatgccaaagaatggcagattctctgtgtccagagatg 2340
QY 2341 ATTCTCGAAGAGGTGAAGATCATGTGAGGCTCAAGGACCCCAAGATCATTTGGGCTGC 2400
Db 2341 attctcgaagaggtgaagatcatgtgaggtcgaagaccaccaatacttctgtctgc 2400
QY 2401 TGGGCGTGTGTGTCAGAGAGACCCCTCTGTGACATGATTACTGACTACATGGAGACGGC 2460
Db 2401 tgggctgtgtgtcagagagacccctctgtgacatgattactgactacatgagaaaggcg 2460
```

QY 2461 ACCTCAACAGTTCCTCACTGCCCCACACACTGAGAGACAAGCAGCCGAGGGGCCCTG 2520
 |||||
 Db 2461 accccaacagctcctcaagtccacacagctgagagacaaagcagccggggccctg 2520
 |||||
 QY 2521 GGGAGCGGAGGCTGGCGAGGGGCCACCATGAGCTAACCAATGCGTGTGATGAGGAG 2580
 |||||
 Db 2521 gggagcggagagctgcgcagggccacacacagccaaatgcgtgcatactgagcag 2580
 |||||
 QY 2581 CCCAGATCGCTCCGGCATGCGCTATCTGCGCACACTCACTTGTGATCGGAGCTG 2640
 |||||
 Db 2581 cccaagctcccgcatgcatctatctgcacactcaactctgtaactcgtgagacctg 2640
 |||||
 QY 2641 CCACCGGGAACCTGCTAGTGGGGAATTTTACATCAAAAATCCGACACTTTGGCATGA 2700
 |||||
 Db 2641 ccacgcygagactgctagtggggaatctacacatacaaatcgcagactcttgcatga 2700
 |||||
 QY 2701 GCCGGAACCTATGCTGGGGGACTATTACCGGTGACAGGCCGGGCGAGTGCCTCCATCC 2760
 |||||
 Db 2701 gccggaacctatgctctggggaactatcagctgagggccggccagctgcgcatcc 2760
 |||||
 QY 2761 GCTGATGCGCTGAGAGTGCATCTCATGAGGAGATTCACGACTCGAGTGCATGTGGG 2820
 |||||
 Db 2761 gctgagatgctcgtgagagatccctcatgaggaggttcaagctcgtgagcgtgag 2820
 |||||
 QY 2821 CCTTGTGTGACCTGTGAGGAGGTGCTGATGCTGTGAGGGCCAGCCCTTTGGGAGC 2880
 |||||
 Db 2821 ccttgggtgacccctgtggaggtgctgagtctctgtagggccagcccttgggagc 2880
 |||||
 QY 2881 TCACGAGAGAGGTGCATCGAAGCGGGGGAGTTCTCCGGGACCAAGGCCCGGACAG 2940
 |||||
 Db 2881 tcacgagagaggtgcatacgagaaagcgaggagttctccgggacagggccggagag 2940
 |||||
 QY 2941 TGTACCTGCTCCGCGGCGCTGCTGCGCCAGGGGCTATGAGTGTGATGCTGCTGCT 3000
 |||||
 Db 2941 tgtacctgtcccgcgccgtcctgcgcagggcctatagagctgtgtctgctgagct 3000
 |||||
 QY 3001 GGAGCGGGAGGTGTGAGAGAGGACCAACCTTTCCAGACTGCATGCGTGTGCGAGAG 3060
 |||||
 Db 3001 ggagcggaggtgtgagagagagacacaccccttccagctgcatgctcctgagagag 3060
 |||||
 QY 3061 ATGCACTCAACAGGTTGATGATCAACATCGAGTGGCCCTCCCTCAGAGGATGATCCAG 3120
 |||||
 Db 3061 atgcaactcaacaggtgtgatacaacatccagctgccccctccagggagatccag 3120
 |||||
 QY 3121 GGGAGCCAGTGTGACATAAACAAGAGACACAATGCGACCTGTGCTCCCTCCCGA 3180
 |||||
 Db 3121 gggagccagtgacataaacaagagacacatgacactctgccttccctcccgga 3180
 |||||
 QY 3181 CAGCCCATCAGCTCTAATAGAGCAATGAGACTGAGTGGGCTGGGCCACCCAGGAG 3240
 |||||
 Db 3181 cagcccatcacctctaataagagagagctgagctgagctgggcccacccagggag 3240
 |||||
 QY 3241 CTGATGCCCTTCTCCCTCTCTGAGACACTCATGCTCCCTCTCTCTCTCTCTCTCC 3300
 |||||
 Db 3241 ctgatgcccttctccctctctctgagacactctcaactgctccctctctctctcc 3300
 |||||
 QY 3301 TAGAAGCCCTGTGCGCACCCAGCTGCTGTGATGAGTGGGATCTCTCCACCTCTCT 3360
 |||||
 Db 3301 tagaagccctgtgcgcacacccagctgctgtgagatgattctctccactctctct 3360
 |||||
 QY 3361 AGCCATCCCTTGGGGAAGGTGGGGAATATAGATAGACACTGAGATGGCCCATTTG 3420
 |||||
 Db 3361 agccatcccttggggaaaggtggggaaatataagataagataagataagataag 3420
 |||||
 QY 3421 GAGCAGCTGGGCGCCACTGGACAACACTGATTCCTGAGAGGTGGCTGCGCCAGCTTC 3480
 |||||
 Db 3421 gagcagctgggcccactggagacaactgattctctgagaggtgctgccccagcttc 3480
 |||||
 QY 3481 TCTGCTCCTGTCAACACTGAGCCCACTGCTGAGAAATGTGGGGGTGAGAGAGCAAGA 3540
 |||||
 Db 3481 tctctcctgtcaacactgagcccaactgctgagaaatctgggggtgagagagacaaga 3540
 |||||

QY 3541 AGGAGAGAAAAATGTTTCTCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600
 |||||
 Db 3541 aggagagaaaaatgtttctctgtcctgctgctgctgctgctgctgctgctgctgctg 3600
 |||||
 QY 3601 CTCCTTCATCAGCTGAAACACTGAGACTTGGGGTAAAGCCGCCAGCCCTCAGTACCC 3660
 |||||
 Db 3601 ctctctcatcacctgaacaactgagactggggtagccccgcgccagccctcagtaacc 3660
 |||||
 QY 3661 CCAGTTCCTCCAGTTCGAGTCTGTAGTGAACCTCTCTTAAGCTTCTGAGGAG 3720
 |||||
 Db 3661 ccacttccactgcaagctgtgtagtaagaaactctcctaagcttaagcttctgtgag 3720
 |||||
 QY 3721 TAAATATTGGGATTTGGGGAAGAGAGAGCAAGCGCCATGAGCTTGGGGTTGGACATC 3780
 |||||
 Db 3721 taatatattggattgggggaaagagagcaagccatgactcttgggttggatcac 3780
 |||||
 QY 3781 TCTAGTGTAGCTGCCAATGATTTTCTAATATACATTTGGGTTGTACATTTTGGG 3840
 |||||
 Db 3781 tctagttagctggccaatgatttttctataatcacttgggttgtacatlttggg 3840
 |||||
 QY 3841 GGAGAGACAGATTTTACACTAATATATAGACTAGCTTACGCAATTTATCCCT 3900
 |||||
 Db 3841 ggaagacacagatttctacataataatagcctagcttgaagcaatttataccct 3900
 |||||
 QY 3901 GCAGTGGCAGGTAATTAATTAAGTTGATTTCACAAAAAATAAATAAACCAGAT 3960
 |||||
 Db 3901 gcactaggcaggtataataaagttgagtttccacaaaaaataaataaaccggaat 3960
 |||||
 QY 3961 TC 3962
 |||||
 Db 3961 tc 3962

RESULT 4

AA084782 standard; DNA; 3754 BP.

AA084782;

17-AUG-1995 (first entry)

Protein-tyrosine-kinase PTK22.

Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;

breast tumor; mamma carcinoma; diagnosis; prognosis; therapy; ds.

Homo sapiens.

Key Location/Qualifiers

FT CDS 142..2886

FT CDS /*tag- a

W09502187-A.

19-JAN-1995.

08-JUL-1994; 94MO-GB01480.

09-JUL-1993; 93GB-0014271.

(CANC-) CANCER RES INST.

(WELL) WELLCOME FOUND LTD.

Barker KT, Crompton MR, Gusterson BA, Martindale JE;

Mitchell PJ, Page MJ, Spence P;

P-PSDB; AAR71100.

PFI: 1995-066991/09.

Method for screening substances, using protein tyrosine kinase

for potential utility as therapeutic agents for cancer

Claim 1; Page 26-30; 51pp; English.

XX cDNA derived from tumor metastatic tissue was amplified using
CC primers (given in AA084783-84) based on sequences (AA71101, AA71103)
CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was
CC identified in an isolated subclone. The 3' sequence of PTK22 was
CC obtained by reverse transcription (using the primer of AA084786) and
CC PCR amplification (primers AA084787-88) of RNA of human breast
CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22
CC is given in AA084782.

XX Sequence 3754 BP: 713 A; 1145 C; 1121 G; 775 T; 0 other;

1
Query Match 91.3%; Score 3617.6; DB 16; Length 3754;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3728; Conservative 0; Mismatches 19; Indels 29; Gaps 7;

QY 181 CCCGGTCGAGCGCTGGGTCTGCCGGAGAGCGATGAGAGTGTGAAGTGGCTAT 240
DB 2 ccggctcgaccgcctgctcgcgggaagagcagatgaggtgtctggaagtggtcat 61
QY 241 TCACTGAGCGATGGGGTGGATGAGGATGCCAAGATGCTGCCACCCCTTA 300
DB 62 tcactgagcgatggggtgtgacttgaaagaaatgcacaaagatgtgtcccccacccctta 121
QY 301 GCGCCGAGGATCAGAGCATATGAGACAGGCGCTGTCCTTTACTGCTGCTCT 360
DB 122 ggcgcgagagatcagagcatatgagaccagagagccgttcccttcttactgtctgtcct 181
QY 361 TGGTGGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 182 tgggtgcaatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 241
QY 421 CCGTGGGCAATGAGACAGGACCATCCCAAGATGATGATGATGATGATGATGATG 480
DB 242 cccgtggcaatgagacagagacacacacacacacacacacacacacacacacacac 301
QY 481 CAGATTCATGTCG 540
DB 302 cagatccacatcgccgc 361
QY 541 GCCCGCAGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
DB 362 gcccgagaggtcggt 421
QY 601 TCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 422 tccactggt 481
QY 661 TCTCCGAGGATACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 482 tctccgagagatcagcg 541
QY 721 ACCGCTGGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 542 accgctgggagagagatgagatgagatgagatgagatgagatgagatgagatgagatg 601
QY 781 ACCCTGGGCG 840
DB 602 accctgggcccccaatgt 661
QY 841 TGAGT 900
DB 662 tgagtggt 721
QY 901 ACACCGCCCTGTGGGAGAGCAATGATTTATCTGAGCGCTGATACCTCAAGCATCA 960
DB 722 acacgccccgt 781
QY 961 CCTATGAGGAGCATACCTGGGCGGAGCTGAGATGAGGAGTGTGGCGCAGCTGGCAGATG 1020
DB 782 cctatgagagacataccgt 841

QY 1021 GTGTGTGGGCGCTGCATGATCTTTAGAGAGATGTCAGAGAGCTGGGGTGTGGCAGGCTATG 1080
DB 842 gttgt 901
QY 1081 ACTATGTGGAGTGGAGGCAACACAGCTTCTCCAGTGGCTATGTGAGATGAGTTGAGT 1140
DB 902 actatgtggagtggaggaac 961
QY 1141 TTGACCGGCTGGGCGCTTCCAGGCTATGAGTGGCTATGAGTGGCTATGAGTGGCT 1200
DB 962 ttgaccggt 1021
QY 1201 GAGCGCTGTGCGCTGGGCGGAGTGTGCTGCTGCGGCTGCGGCTGCTGCGGCTGCGG 1260
DB 1022 gagcgt 1081
QY 1261 AGGGGAGCCCATGCGGCAACACTGAGGGGCAACCTGGGGAGCCCAAGACCGGGCTG 1320
DB 1082 aggggagcccatgctgcac 1141
QY 1321 TCTCAGTCCCGCTTGGGCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 1142 tctcagtgccccctgt 1201
QY 1381 GCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1202 ggcctgt 1261
QY 1441 CGGCACTGGGAGGACCTTCCCGGAGCCCGCTGTGGGCGGCTGCGGCTGCGGCTGCG 1500
DB 1262 cggcactgggaggaacttcccgagcccgccctgtgtgtgtgtgtgtgtgtgtgtgtgt 1321
QY 1501 ACTTCAGAGCTTGGAGCTGAGAGCCCAAGAGCCCAAGC---AGCCGCTGCAAGGCGAGG 1557
DB 1322 acttcagagcttggagctggagctggagctggagctggagctggagctggagctggagct 1381
QY 1558 GGAAGCCGAGCCGATCCTGATCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1617
DB 1382 ggaagccgagagcgcacatctcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1441
QY 1618 TCAATGCGCTCATGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1677
DB 1442 tcaatgcccccatgt 1501
QY 1678 GGGGTGTGGAGAGAGAGCTACAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1737
DB 1502 ggggtgtggagagagagctgacagcttccacacacacacacacacacacacacacac 1561
QY 1738 ACAAGCGCCGAGGCTCTAGAGAGCAACCCCGTACAGAGAGCCCGGCTGCTGAGATC 1797
DB 1562 acaagccgcaggtctctag 1621
QY 1798 CGGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1857
DB 1622 cggcgcaactgcgtctctgt 1681
QY 1858 ACCGCTCTCTTCTGCGGCACTTACGCGCTGCTGCGGCTGCTGCGGCTGCTGCGG 1917
DB 1682 accgctctcttctgagcaatcagccccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1741
QY 1918 CTTGGGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1977
DB 1742 cttgggcccccaaaccaaaccaaaccaaaccaaaccaaaccaaaccaaaccaaaccaac 1801
QY 1978 CAGGCGCCCGCTTGTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 2037
DB 1802 caggcgccccgttctgtccccacacacacacacacacacacacacacacacacacacac 1861
QY 2038 ACATGTTAACCTGAGAGGCGCTGACCGGGGAGCAACACTATGCTGTGCTGCTGCTG 2097
DB 1862 acatgttaacctgagagcgctacacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1921
QY 2098 CAGGGGAGTGGGAGATGAGGCGCCCGCAGAGTGGATTTCCCTGAGTGGATCCGCTTCA 2157

PD 21-SEP-2000.
XX
XX 08-MAR-2000; 2000MO-US05989.
PF
XX
PR 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Ruben SM;
XX
XX WPI, 2000-579444/54.
DR P-PSDB; AAB54286.
XX
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing, treating, or ameliorating a medical condition, particular pancreatic cancer, or for use in assays for diagnosing a pathological condition -
PS Claim 1; Page 714-715; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 2861 BP; 567 A; 899 C; 800 G; 589 T; 6 other.

Query Match 70.3%; Score 2783.6; DB 21; Length 2861;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2833; Conservative 5; Mismatches 4; Indels 20; Gaps 3;

QY 1093 GGAGCAACCAAGCTTCCAGTGGCTATGAGATGAGTTGACGGCTGA 1152
DB 1 GGCACATCAAGCTCCAGTGCATGAGATGAGTTGACGGCTGA 60
QY 1153 GGGGCTTCAGGCTATGAGGCTCACTACACATGACACAGGCTGGAGCCCTGCG 1212
DB 61 GGGCTCCAGGCTATGAGGCTCACTACACATGACACAGGCTGGAGCCCTGCG 120
QY 1213 CTGGCGGGGTGGAATGCGCTTCCGCGGTGGCTGCATGAGGCTGGAGGAGCCCA 1272
DB 121 CTGGCGGGGTGGAATGCGCTTCCGCGGTGGCTGCATGAGGCTGGAGGAGCCCA 180
QY 1273 TGGCCCAACACCTAGGAGGAGCACTGGGAGCCCAAGCCGAGCTTCACAGTCC 1332
DB 181 TGGCCCAACACCTAGGAGGAGCACTGGGAGCCCAAGCCGAGCTTCACAGTCC 240
QY 1333 TTGGCGGCTGAGGCTTCTTCAGTGGCGCTTCCTTGGCGGAGCCCTGCTTAC 1392
DB 241 TTGGCGGCTGAGGCTTCTTCAGTGGCGCTTCCTTGGCGGAGCCCTGCTTAC 300
QY 1393 TCTTCAGCGCAATCTCTTCATCTGATGTGTAACAATTCCTTCGCGCACTGGAG 1452
DB 301 TCTTCAGCGCAATCTCTTCATCTGATGTGTAACAATTCCTTCGCGCACTGGAG 360

QY 1453 GCACCTTCCGCGCAGCCCGCTGTGGCGCCTGGGCCCACTCCCAACCACTTACAGACT 1512
DB 361 GACCTTCGCGCAGCCCGCTGTGGCGCCTGGGCCCACTCCCAACCACTTACAGACT 420
QY 1513 TGGAGCTGGAACCCAGAGGCCAGACCCGCTGGCCAAAGGCCGAGGGAGCCGACCA 1572
DB 421 TGGAGCTGGAACCCAGAGGCCAGACCCGCTGGCCAAAGGCCGAGGGAGCCGACCA 480
QY 1573 TCCATCATGCGCTGCTGTGGCCATCATCCGCTCCGCTGCTATCATTTGCTCATGC 1632
DB 481 TCCATCATGCGCTGCTGTGGCCATCATCCGCTCCGCTGCTATCATTTGCTCATGC 540
QY 1633 TCTGGCGGCTGCACTGGCGAGGCTCTCTCAAGAGGCTGAACGAGGCTGTGGAAGG 1692
DB 541 TCTGGCGGCTGCACTGGCGAGGCTCTCTCAAGAGGCTGAACGAGGCTGTGGAAGG 600
QY 1693 AGCTGACGCTTCACTCTCTCTCTGGGAGACATATCTCATCAACACCGCCAGGTC 1752
DB 601 AGCTGACGCTTCACTCTCTCTCTGGGAGACATATCTCATCAACACCGCCAGGTC 660
QY 1753 CTGAGAGGCAACCCCGCTTACAGAGGCGCCGCTGCGGAATCGCCCACTCGCTC 1812
DB 661 CTGAGAGGCAACCCCGCTTACAGAGGCGCCGCTGCGGAATCGCCCACTCGCTC 720
QY 1813 CCTGTGTCCCAATGAGCTGTGCTGTGCTCAATCCAGCTACCGCTCTCTCTG 1872
DB 721 CCTGTGTCCCAATGAGCTGTGCTGTGCTCAATCCAGCTACCGCTCTCTCTG 780
QY 1873 CCACTTACGCGCTCTCTCTCTGAGGCGCCGCGCCCAACCGCTGAGCAACCA 1932
DB 781 CCACTTACGCGCTCTCTCTCTGAGGCGCCGCGCCCAACCGCTGAGCAACCA 840
QY 1933 CCAACACCCAGGCTCACTAGTGGGAGTATATGAGCTGGAAGCCAGCCCGCTTC 1992
DB 841 CCAACACCCAGGCTCACTAGTGGGAGTATATGAGCTGGAAGCCAGCCCGCTTC 900
QY 1993 TGGCCCACTCCCAAGAGGCTGCGCTGCTGCTCAATCCAGGCTGCTTACCTTC 2052
DB 901 TGGCCCACTCCCAAGAGGCTGCGCTGCTGCTCAATCCAGGCTGCTTACCTTC 960
QY 2053 AGGCGCTACCGGGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2112
DB 961 AGGCGCTACCGGGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 2113 ATGGGCGCCCGAGAGTGGATTTCCGATCTGCACTCCGCTTCAAGAGGCTTGGCG 2172
DB 1021 ATGGGCGCCCGAGAGTGGATTTCCGATCTGCACTCCGCTTCAAGAGGCTTGGCG 1080
QY 2173 AGGCGCACTTGGGAGGAGGACGCTGTGAGTGTGAGAGCCCTCAAGATCTGCTGCT 2232
DB 1081 AGGCGCACTTGGGAGGAGGACGCTGTGAGTGTGAGAGCCCTCAAGATCTGCTGCT 1140
QY 2233 TTGATTTCCCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2292
DB 1141 TTGATTTCCCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 2293 GGGCGAGTCAACCAAGAAATGCACTTCTCTTCTTCCAGAGATATTTCTGAAAG 2352
DB 1201 GGGCGAGTCAACCAAGAAATGCACTTCTCTTCTTCCAGAGATATTTCTGAAAG 1242
QY 2353 AGGTGAATGATGTCGAGGCTCAAGAGCCCAACATTCGCTGTGGCGCTGTG 2412
DB 1243 AGGTGAATGATGTCGAGGCTCAAGAGCCCAACATTCGCTGTGGCGCTGTG 1302
QY 2413 TGAAGAGGACCCCTCTGATGATTTACTGATGATGATGATGATGATGATGATGAT 2472
DB 1303 TGAAGAGGACCCCTCTGATGATTTACTGATGATGATGATGATGATGATGATGAT 1362
QY 2473 TCCATCATGCGCTGCTGTGGCCATCATCCGCTCCGCTGCTATCATTTGCTCATGC 2532
DB 1363 TCCATCATGCGCTGCTGTGGCCATCATCCGCTCCGCTGCTATCATTTGCTCATGC 1422
QY 2533 CTGCGAGGGGCGCCACCATGCTACCAATGCTGCTGCTGCTGCTGCTGCTGCTG 2592

|||||
Db 1423 ctgagcggggccacacataagctacacaaatgctgcatgctgagcccaagctgcct 1482
Oy 2593 CCGGATGCGCTATCTGGCCACACTCACTTGTATCGGAGACTGCGCAGCGGAGACT 2652
Db 1483 ccggacatgcctatctgcccacacacacttctacatcgagactgcccagcgagact 1542
Oy 2653 GCGTAGTGGGGAAAATTTCACATCAAAATCGCAGACTTGTGCATGAGCGGAACTCT 2712
Db 1543 gccatgttgggaaaatttcacacatacacaatcgagacttggcatgagccggaacctc 1602
Oy 2713 ATGTGGGAGACTATTACCGTGTGAGGGCGGAGAGTGCCTCCATCCGCTGATGGCCT 2772
Db 1603 atgctgggactattacacgtgtgcaaggcgagctgcctccatcgctgagtgacct 1662
Oy 2773 GGGAGTCATCCATGATGGGGAAGTTACGACTGCGAGTGAAGTGGGCTTGGTGTGA 2832
Db 1663 gggagtcataccctcaatgaggaagtcacagacgcagagtgagtgctgctgctgta 1722
Oy 2833 CCTGTGGAGTGCTGATGCTGTGTAGGGCCAGCCCTTGGGACGCTCACGAGCAGC 2892
Db 1723 cccgtgggaggtgtgtatgtctgtatggccagcccttggcagctcaacgagagc 1782
Oy 2893 AGGTATCGAGAACCGGGGGAGTTCTTCGGGACAGGGCGGAGGTGACTGTCC 2952
Db 1783 agtcatcgagaaacgagcgaggagttcttcgagacagcgagcgagtgatcctgccc 1842
Oy 2953 GCGCGCTGCTGCTGCGCCAGGCGCTATATGAGCTGATGCTGCTGGTGGAGCGCGGAGT 3012
Db 1843 ggcgcgcctgctgcccga .ggcyatatgtgactgactcgtgctgctgagcgagtg 1901
Oy 3013 CTGAGCAGCGCACCCCTTTCACAGTGCATCGGTTCTGTGGCAGAGGATGACTACACA 3072
Db 1902 ctgagcagcgacacaccccttcccaagctgcatgctgctgctgagagatgactcaaca 1961
Oy 3073 CGGTGTGAATCACACATTCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAGCCAGTG 3132
Db 1962 cgggtgatacacacatccagctgcccctccctcagagagatccagggagccagctg 2021
Oy 3133 ACATTAACAAAGAGAGACACAATGGCAGCTGCGCTTCCCTCCGACACCCATCACC 3192
Db 2022 aactataaacaagagagacaatgacactgtccctccctccctccgacagccataccc 2081
Oy 3193 TCTATATAGGACAGTGAAGTGCAGTGGGTGGGCCACCGAGAGACTGATGCCCTT 3252
Db 2082 tctaataagagcagtgtgacacgagtgagctgagcccaacagggagctgtagccctt 2141
Oy 3253 CTCCCTTCTGGAACACTCTCATGTCCTTCCCTGTTCTTCTCTAGAACCCCTG 3312
Db 2142 ctcccttccctggacacactctcatgtcccttctcttctctctagaaagccctg 2201
Oy 3313 TCGCCACCCAGCGTGTCTGTGATGGGATGCTCTCCACCTCTCTACGATCCCTTG 3372
Db 2202 tcgcccacacagctgtgctctgtgaltgagatctctccactctctctcaagcatccctg 2261
Oy 3373 GGGAAAGGTGGGGAATATATAGATATAGACTGACATGAGCCCATTTGAGACCTGGGC 3432
Db 2262 gggaaaggtgggaaataatagatagacatgagacatgagccatgtagcacttgagc 2321
Oy 3433 CCCACTGGAACAACACTGATTCCTGTAGAGAGTGGCTGCG -CCCGAGCTTCTCTCCCTGT 3491
Db 2322 cccactggaacaacactgttctctgtgaggtgctgctgcccacagcttctctctcctgt 2381
Oy 3492 CACACACAGGACCCACACGCTGCTGAGAAATCTGGGGTGGAGGAGAAAGAGAGGAAA 3551
Db 2382 cacacacagggacccacacgctgtgagaaatctgggggtgaggaagaagaaggagaaa 2441
Oy 3552 ATGTTCCTTGTGCTGCTCTGTACTGTCTCAGTGTGGGCTTCTCTCTCTCATCA 3611
Db 2442 atgttctctgtgctgtctctgtactctgtctctcagcttggcttctctctctcatca 2501
Oy 3612 CCTGAACACTGAGACTGGGGGTAGCCCCGCCACAGCCCTCATGTACCCCACTTCCAC 3671
|||||

Db 2502 cctgaacactgacacctgggtgtagcccgcccgccctcagctacacccactccac 2561
Oy 3672 TTGCAGTCTTGTAGCTAGAACTTCTTAAGCTTATACGTTTCTGTGAGTAATATTGGG 3731
Db 2562 ttgcagcttctgtagcactctctcagcctataagcttctgtgagtaaatatttggg 2621
Oy 3732 ATTGGGGGAAAGAGGAGAGCAAGGCCATATGCTTGGGGTGGACATCTGTAGTAGC 3791
Db 2622 attgggggaaagagagagcaagccataagccttgagggttgagacatctcagtgagc 2681
Oy 3792 TGCCACATTTGATTTTCTATATATCACTTGGGGTTGTACATTTTGGGGGAGAGACACA 3851
Db 2682 tgccacatatttctctataatcacttggtgttgatcatatttgaggagagacaca 2741
Oy 3852 GATTTTACACTAATATATAGCACTTACCTTGAAGCAATTTATCCCTGCACTAGCAG 3911
Db 2742 gattttacactaataatagctagcttgaagcaatttaacccctgactagagag 2801
Oy 3912 GTAAATAATAAGTTGAGTTTCCACAAAAAATAAAAAA 3953
Db 2802 gtaataataaggttgaatttccacaaaaaataaaaaa 2843

RESULT 6
AAH3198
ID AAH3198 standard; cDNA: 2861 BP.
XX
AC AAH3198:
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:254.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 6; ss.
XX
OS Homo sapiens.
XX
PN M020012920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WP1: 2001-235357/24.
DR P-PSDB: AAG73767.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 1; Page 2388-2389; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204

CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX

SO Sequence 2861 BP; 567 A; 899 C; 800 G; 589 T; 6 other;

Query Match 70.3%; Score 2783.6; DB 22; Length 2861;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2833; Conservative 5; Mismatches 4; Indels 20; Gaps 3;

QY 1093 GGAGCAACACAGCTTCTCCAGTGTGAGATGAGATTGATGACGGCTGA 1152
D 1 ggcacacacacgcttcccgagtgatgtgagatgagattgagtttgacgagcga 60
QY 1153 GGGCTTCACAGGCTATGACAGGTCCACTGTAAACATGACACAGCTGGAGCCGCTGAC 1212
D 61 gggcttcacagctatgacagtgacatgtaaacacacgacacgctggagccgctgc 120
QY 1213 CTGGCGGGGAGGAAATGCGCTTCCGGCGCTGGCGCTGCGCATGGCTGGAGGGAGCCCA 1272
D 121 ctggcggggaggaatgctgcttccggcgctggcgctgacatggctggagggagccca 180
QY 1273 TGGCCACACACCTAGGGGGCAACCTGGGGGACCCAGAGCCGGGCTGTCACAGCCCC 1332
D 181 tggccacacacctagggggcaacctgggggagcccaagccgggctgtctcagtgcccc 240
QY 1333 TTGGCGGCGCTGTGGCTTCTGCAAGTGCCTTCTTGGCGGGGCGCTGTTAC 1392
D 241 ttggcgggcgctgtggcttctgcaagtgccttcttggcgggcgctgtttac 300
QY 1393 TCTTACAGCAAAATCTCTTCAATCTCTATGTGTGAACAAATCTCTTCGGCACTGGAG 1452
D 301 tcttacagcaaatctcttcatatcttgatgtgtgaacaattctctcggcactggag 360
QY 1453 GCACCTTCCCGCAGCCCGCTGGTGGCGGCTGGCCACCTCCCAACCACTTCAGACACT 1512
D 361 gcaaccttcccgacgcccgcctgggctggcgctggccacactccacaacacttcagact 420
QY 1513 TGGAGCTGAGCCAGAGGCCAGAGCCCGTGGCCAAAGGCGAGGGAGCCGACCGCA 1572
D 421 tggagctgagccagagcccgctgggctggcgctggccaaaggcgagggagcccgacca 480
QY 1573 TCCCTCAATCGGCTGCTGTGTGTCATATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
D 481 tccctcaatcggtctgt 540
QY 1633 TCTGCGGGGCTGCACTGAGCGGAGGCTCTCAAGAAAGGCTGAACGAGAGGTGTGAAGAG 1692
D 541 tctgcggggctgcaactgagcgaggttctcagcaaggtgaaacgagaggtgtgtgaagag 600
QY 1693 AGCTGACGGTTCACTCTGTCTGTCCTGGGACACTATCTCATCAACAACCGCCAGATC 1752
D 601 agctgacggttcacctctgtctgtcccgaggacatactccatcaacaacgcccaggttc 660
QY 1753 CTAGAGAGCACCACCCCGTACAGAGACCCCGGCTGTGGGAAATCCGCCACTCCGCTC 1812
D 661 cttagagagcaccacccgttacagagagcccgagctgtgggaattccgcccacatcgtctc 720
QY 1813 CCTGTATCCCAATGAGCTGTGCTGTGCTGCTCAATCCAGCTACCGCTCTTGTGG 1872
D 721 cctgtatcccaatgagctgtgctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
QY 1873 CCACTTACGCGCGCTCCCTCGAGGCGCGGCGCCCAACACCGGCTGGGCAAAACCA 1932
D 781 ccaacttacgcgctccctcgtgagcgcgcgcccccacacacgctgtggcacaaccca 840
QY 1933 CCAACACCGAGGCTTACAGTGGGAGTATATGAGCCTGAGAAAGCAGGCGCGCGCTTC 1992
D 841 ccaacacccgaggttaccagtgaggacatatagtgagcctgagaaagcagcgcccgcttc 900

QY 1993 TGCCCCCACTCCCCAGAACAGCGTCCCCATTATGCGAGGCTGACATGTTACCTGC 2052
D 901 tcccccaactccccagaaacagcgtcccatatgtccgaggtgtgacatgttaacctgc 960
QY 2053 AGGCGCTACCGGGGGCAACACTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2112
D 961 aggcgttacccgggggcaacactatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
QY 2113 ATGGGCCCCCAGAGTGAATTTCCCTCGATCTGCACTCCGCTTCAAGAGAGGCTTGGG 2172
D 1021 atgggcccccaagatgt 1080
QY 2173 AGGCGCAGTTTGGGAGTGTGCACTGTGTGAGGTGACAGCCCTCAAGATCTGTGATGC 2232
D 1081 aggcgcagtttgggagtgacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
QY 2233 TTGATTTCCCTTATGTGCTTAAGGAGACACCTTTGTGTGATCTGCAAGATCTTAC 2292
D 1141 tggatttcccttattgtgcttaaggagacaccttctgtgtgtgtgtgtgtgtgtgtgtgt 1200
QY 2293 GGCAGATGCCACCAAGATGCCACTTCTCTTGTCTCCAGGAATGATTTCCAGAAAG 2352
D 1201 ggcagatgcccacaaagatg-----ccaggaaatgtatcttcgaag 1242
QY 2353 AGGTGAAGATCATGTGTGAGGCTCAAGAACCCCAACATCATTCGCTGGCGCTGTGTG 2412
D 1243 aggtgaagatcatgtgtgaggtgtcaaggaacccaatcatctgtgtgtgtgtgtgtgt 1302
QY 2413 TGCAGAGAGACCCCTCTGCAATGATTAATGACTACATGAGAGAACGGCACTCAACAGT 2472
D 1303 tgcagagagacccctctgcatgtatctgactatgacatgtgagaaagcgacctcaacag 1362
QY 2473 TCTCAGTGGCCACACAGCTGGAGGAGCAAGGCAAGGAGGGGCGCTTGGGAGCGGAGG 2532
D 1363 tctcagtgcccaacagctgt 1422
QY 2533 CTGGCAGAGGGCCCAACATCAGTACCCCAATGCTGCTCAATGTGGACCCCAAGATCGCT 2592
D 1423 ctggcagagggcccaacatcagctacccaatgctgtgtgtgtgtgtgtgtgtgtgtgtgt 1482
QY 2593 CCGGCTATCGCTATCTGTGGCCACACTCAACTTGTGTACATCGGAGCTGGCAAGCGGA 2652
D 1483 cgggctatcgctatctgt 1542
QY 2653 GCCTAGTGGGGAATTTCAACATCAAAATCGCAAGACTTTGGGATGAGCGGAACTCT 2712
D 1543 gctcagtggtggaaatcttccacatcaaaatcgaagacttggatgtgacggaaactct 1602
QY 2713 ATGCTGGGACTATTTACCGTGTGAGGCGCGGAGTGTGCTCCATCGCTGATGGGCT 2772
D 1603 atgctgggactatttacgt 1662
QY 2773 GGGAGTGCATCTTATGAGGGAAGTTCAACAGACTGCAAGTGAAGTGTGGGCTTGGTGTGA 2832
D 1663 gggagtgcattcttaagt 1722
QY 2833 CCTGTGGGAGGTGCTGATGCTGTGTGAGGCGCGGCTTTGGGCACTACCGACAGC 2892
D 1723 cctgtgggaggt 1782
QY 2893 AGGTATCGAAGAACCGGGGAGTTCTTCGGGACACAGGCGCGGACAGTGTATCTGTGCC 2952
D 1783 aggtatcgaaagacggt 1842
QY 2953 GGGCGGCTGCTGTGCGCAGGCGCTATATAGCTGATGCTTCCGTGTGAGGCGGGAGT 3012
D 1843 gggcggt 1901
QY 3013 CTGAGCAGCAGCAACCTTTTCCAGCTGATGAGGCTTCCGTGAGAGAGATGCACTCAACA 3072
D 1902 ctgagcagcagcacccttccagctgtcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1961
QY 3073 CGGTGTGAATCAACATCAAGCTGCCCTCCCTCAAGGAGTGTATCCAGGGAAGCCAGTG 3132

[illegible]

RESULT	10
AAT93784	
ID	AAT93784 standard; cDNA; 3157 BP.
XX	

AC	AA193784;
XX	
DT	16-FEB-1998 (first entry)
DE	
XX	CKK-2, a human mammary carcinoma kinase 10 MCK-10) family member, cDNA.
XX	
KW	Mammary carcinoma kinase: MCK-10; CKK-2; receptor tyrosine kinase;
KW	proliferative disease; cancer; MCK-10 activity; aberrant expression; ds.
XX	
OS	Homo sapiens.
XX	
FX	
FX	Key Location/Qualifiers
FT	370..2934
FT	CDS
FT	/*tag- a
XX	
PN	US5677144-A.
XX	
PD	14-OCT-1997.
XX	
PF	08-NOV-1994; 94US-0336343.
XX	
PR	16-NOV-1993; 93US-0153397.
XX	
PA	(ALVE/) ALVES F H E.
PA	(ULRL/) ULLRICH A.
XX	
PI	Alves FHE, Ullrich A;
XX	
DR	WPI: 1997-511869/47.
DR	P-PSDB; AAM34671.
XX	
PT	Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding
PT	for it, useful for cancer diagnosis
XX	
PS	Claim 9; Fig 3; 70pp; English.
XX	
CC	The present sequence represents the cDNA sequence of human CKK-2, a
CC	member of the mammary carcinoma kinase 10 (MCK-10, AA193785) family of
CC	receptor tyrosine kinases. Expression of CKK-2 is associated with
CC	proliferative diseases such as cancer. The CKK-2 gene was identified by
CC	PCR and a cDNA prepared from colonic adenocarcinoma RNA. CKK-2 is
CC	expressed in a wide variety of cancer cell lines and tumour tissue. The
CC	CKK-2 nucleic acids can be used for diagnostic purposes to detect
CC	aberrant expression of CKK-2 genes. Engineered cell lines, containing
CC	recombinant vectors with the present sequence, are useful for producing
CC	infectious retroviral particles. The cell lines may also be used to
CC	evaluate and screen drugs involved in CKK-2 activation and regulation.
XX	
SO	Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T; 0 other;

Query Match	16.2%	Score 642	DB 161	Length 3157
Best Local Similarity	56.5%	Pred. No. 5.3e-131		
Matches 1534	Conservative	0	Missmatches 975	Indels 207
			Gaps	10
QY	TCGCTGCTGCTCTGGTGGCAAGTGGAGATGCTGCATGCATGAAGGACATTTTGTGATCCCGCA	408		
Db				
	tggtgtctgtctctgtctgtctgtcctatcttgahtctgcgaagctcctggtatcatccgca	454		
QY	AGTGCCGCTATGCCCTCGGGCATGACGAGACCGGACCATCCAGACATGCATCTGCTT	468		
Db				
	tatgcgcgtatctctctgtgcacatgcattgcaaggcgcagattccagatgagacatcaacgctt	514		
QY	CCAGTCCMGCGCAATTCACATCCAGTCCGCCGCCGCCACAGCAGGTGGAGAGACGTAGACGGG	528		
Db				
	ccagtcagtgvcgacagatccacagctgcacaaatataggaagtgctgactcaagaagaagg99	574		
QY	ATGGGCGCTGTGATGCCGCCACAGGATGGGTGTTTCCCAAGGA--GGAGAGATCTACTGACG	585		
Db				
	atggagcctgtgtcccttgaagattccagtggaacctgagactggaaggaagtttcttcgcaga	634		
QY	TGATCTTCAACAGCATCCACTGCTGTGCTGTGGTGGGACCCAGGAGCGCATGCGCGGG	645		

Dh 635 ttgacttgcacacccctccatcttatccactctgttgggaaccaggggcgcatgcagag 694
Qy 646 GCCTGGGCAAGAGATTCTCCGGAGCTACGGGCTGCTACTCCGGAGTGGTCCGCCCT 705
Dh 695 gtcaatgcacatcgatttgcctccatctgatacaagatcaattacggtgagctgcctcgt 754
Qy 706 GGATGGGCTGGGAAGACCCGCTGGGCTCAGGAGTGTATCTCAGGCAATGAGACCTCGAGG 765
Dh 755 ggaatcttggcgggaaccgttcaatggaaacaggtgtggtgataagtaagtaaccctatg 814
Qy 766 GAGTGTCTGAGAGACCTTGGGCCCCCATGTTGGTCCGAGTGGTGTCTTACCCCC 825
Dh 815 acaatttccataaagaacttggagccgccaattgtacagcaattgtccggttccattccag 874
Qy 826 GGGCTGACCGGGTCAATGAGTGTCTGTGCGGGTAGAGCTCTATGAGCTGCTCTGAGAGG 885
Dh 875 tcaccgacacatcagtaagtgtgtatgagagtgaggtttaaagcgtgtgtctgtctag 934
Qy 886 ATGGACTCTGTCTTACACCCGCCCTGTGGGCAACAATGTATTTATCTGAGG----- 939
Dh 935 atggtcttgggtcttaccatgtctcagctggtgcaagcttgttactccctggaagttcca 994
Qy 940 CCGTATACCTCAAGACCTCCACCTATGAGAGGACATACGTTGGGCGGAGTGAATGAGG 999
Dh 995 tcaattatcgtgaatcctgtctatgtatgagctgttggatcacagcatgacagaagag- 1053
Qy 1000 GTCTGGCCAGCTGGCAGATGCTGTGTGGGCTGAGGCTTCCAGGCTATGAGAGTCAAGAGC 1059
Dh 1054 --ctagagcaatgtacagatgtgtgtctgtcgtgagcaatttccaccagaccatgaat 1111
Qy 1060 TCGGGGTCTGGCCAGCTATGACTATGTGGAGTGAAGACACACAGCTTCTCCAGTGGCT 1119
Dh 1112 accagctgtggtccgctatgtatgtgtggtggtgagagagaggtgcccacatgtgt 1171
Qy 1120 ATGTGAGATGAGATTGATTGATTTGACCGGCTGAGGCTTCCAGGCTATGAGAGTCACT 1179
Dh 1172 acattgagatcaatgttgaatttgaatttgcacatgaatttccatccatgaagttccact 1231
Qy 1180 GTATCAATATGACACGCTGGGAGGCCGCTCTGCTGGGCGGTGGATGATGCTTCCGCG 1239
Dh 1232 gcaataacatcttcttaaggtgtgtaagatctttaaaggatgtaagatgctctccgct 1291
Qy 1240 GTGGCCCTGCAATGAGCTGGGAGGAGGCCCATGCGCACAACTAGAGGGCAACTGG 1299
Dh 1292 ctg---aagccagtgatggtggaactaatgtccatttcccttcccttgcgtgtagc 1348
Qy 1300 GGGACCCCAAGACCCGGGCTGTCTAGTGGCCCTTGGGCGGCTGTGGCTTCTTCTGCT 1359
Dh 1349 tcaaccccaatgtctgtgttctacaggtgtcttccaccacacgaattgtgccaatgca 1408
Qy 1360 AGTGGCTCTCTCTTGTGCGGGGCTGTACTCTTCAAGCAAAATCTCTCATCTGCTG 1419
Dh 1409 agtgcataatacatcttgcagataactctgtaagtgttcaagtgtgagatccacttccaa 1468
Qy 1420 ATGTGTTAACAATCTCTCTCGGCACTGGAGGCACTTCCCGGCAAGCCCTGCTGGG 1479
Dh 1469 atgtgcataatgttacaacaactctgagccctgtccacacttcccta----- 1513
Qy 1480 CGGCTGGGCAACCTCCCAACAACCTCAGAGCTTGGAGCTGGAGGCCCAAGGCCACGAC 1539
Dh 1514 -----tggaaccacaacataatgac 1534
Qy 1540 CCGTGGCCAAAGAGGAGGAGGCCGACCCGACCTCATGCTGAGCTGTGGTGGGCAATCA 1599
Dh 1535 caatgtctaagtgtatgtaagcaaacactcgatcttccatgtgtgtgtgtgtgtgtgtgt 1594
Qy 1600 TCTGCTCTCTGCTGATCATATGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1659
Dh 1595 tcttatctctctgtgcctcttctgtcatcaccctctgtgagcagttctgtgagaaatgc 1654
Qy 1660 TCAGCAAGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1719
Dh 1655 tggagaaggtcttccgtgagagatgtgtgataagtaagtaagtaagtaagtaagtaag 1714

Qy 1720 GGGACATATCTCATCAACACCGCCAGTCTTAGAGACCAACCCCGTACCAGAGAGC 1779
Dh 1715 gtcatttcatcagatgttcaacaat-----aaacgtctct 1747
Qy 1780 CCGGCTCTGTGGGAATCCGCCCAATCCGCTTCCCTGTCTGCTCCCAATGAGCTGTGCTG 1839
Dh 1748 catcacatgtatgaacaaggtgtccaaactcagatcagatgtcatctt----- 1794
Qy 1840 TGTCTTCAATCAAGCTACCGGCTCTTCTGTGGCACTTACGCGCGTCCCGCTCGAGG 1899
Dh 1795 -----cccttcgtcc 1804
Qy 1900 CGGGCCCCCAACACCCGCTGGGCAACCAACCAACCCAGGCTTACAGTGGGAGCT 1959
Dh 1805 ctgactaccagagagcatccatcagctgtatagaaactccagaaattgtctccagggag 1864
Qy 1960 ATATGAGCTTGAAGAACCAAGCGCCGCTTCTGCCCCACCTCCCAAGAACAGCTCC 2019
Dh 1865 aggaatcagatgtcagcgtgtgtgtgaaagccagttccagccagttgcccctgaggggtgc 1924
Qy 2020 CCAATTATGCGAGGCTGACATTTGATTCAGGCGCTACCGGGGCAACACTATG 2079
Dh 1925 ccacatactgagaggtgtgacatagtaacatccaaagagtgacaagagcaacacatct 1984
Qy 2080 CTGTGCTGCACTGCCCCCAGGGGCGAGTC---GGGATGGGCCCCCAGAGTGAATTTCC 2136
Dh 1985 cagtgtcgtcgttccacatagacatgtctctcaggaanaaatgtgtgtgtgtgtgtgtgt 2044
Qy 2137 CTGATCTTGACCTCCGCTTCAAGGAGAGACTTGGCGAGGCGCAATTTGGGAGGTCAC 2196
Dh 2045 ccagaaactcctaacttcaagaagagctgtggaagagagagagagagagagagagagag 2104
Qy 2197 TGTGAGAGTCCACAGCCCTCAAGATCTGTGATGTTTATTTCCCTTATATGTCGTA 2256
Dh 2105 tctgtgaagtgtggaagatgtgaataatcaagaacaagaatttgccttagatgtatgt 2164
Qy 2257 AGGACACCTTGTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2216
Dh 2165 ccaacacagctctcgt 2221
Qy 2317 GCTTCTCTTGTTCACAGGAATGATTTCTGTAAGAGAGTGAAGATGTGAGGCTCA 2376
Dh 2222 -----ccaggaatgtatttcttaaggaatgaatgaatgaatgaatgaatgaatga 2266
Qy 2377 AGGACCCCAACATCATTTGCGGCTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2436
Dh 2267 aggaaccaacaatcacaatcacaatcacaatcacaatcacaatcacaatcacaatcaca 2326
Qy 2437 TTAAGTACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2496
Dh 2327 tcaatgaataatcaggaatgtgaatgtgaatgtgaatgtgaatgtgaatgtgaatgtga 2386
Qy 2497 ACAAGGACCGAGGGGGGCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2556
Dh 2387 atttcttccacagc-----atgaagcagctgtcagtt 2419
Qy 2557 ACCCAATGCTGTGATGTGGAGCCCAAGATGCGCTCCGGCATGCGCTATCTGGCCACAC 2616
Dh 2420 acaacaatctgaattatgtgtaacaaatgtcctgtgcaatgaatgaatgaatgaatga 2479
Qy 2617 TCAACTTTGATCATGGGAGCTGCGCACGCGGAGCACTTAATTGGGAAATTTACCA 2676
Dh 2480 ttaatttgttcaacgagatctgtgcacacgaactgtttagtggtaagaatacaaca 2539
Qy 2677 TCAAAATGACAGCTTTGGCATGAGCCGAGCACTTATGCTGGGAGCTTACCGTGTGC 2736
Dh 2540 tcaagaatgtctgtatgtgaatgagcagaacctgtacagtgatgtacattaccgagatcc 2599
Qy 2737 AGGCGGGGAGGAGTGTGCGCATCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2796
Dh 2600 agggcgggaggtcttccctacatccatccatccatccatccatccatccatccatccatcc 2659

OY	2487	CAGCTGGGAGCAAGGAGGCCCGCCCTGGGGACGGGACAGGTTCGCAGGGGCC	2546
Dd	1940	gagccccaattcttcctcagcg-----atgtacgc	1972
OY	2547	ACCATCAGCTACCCACTGCTGTGCATGTGGCAGCCAGATGCCCTCGCATGCGTAT	2606
Dd	1973	actgycagttcaccaaatctgaagttaigtgtaccacaattgcctctgcatgaagtac	2032
OY	2607	CTGGGCACACTCACATTGTGTACATCGGAGACTTGCCACCGGAACTCCTTAGTTGGGAA	2666
Dd	2033	ccttcctctctaatttggttcacagatctgcccacacgaacctgttagtggtagaag	2092
OY	2667	AATTTCACATCAATAACGGACAGCTTTGGCATGAGCCGGAACCTGTATGCTGGGACTAT	2726
Dd	2093	aactacacaatacaagaatgaagtctcttggaaatgagcaagacctgtacagttggtacat	2152
OY	2727	TACCGTGTGCAGGGCCGGGACGTGTCGCCATCCGCTGATGGCTGGAGATGCATCTTC	2786
Dd	2153	taccgatccagggccgggagtgctccctatccgtgtgatatgtcttgggagagtacttg	2212
OY	2787	ATGGGGAAGTTCACGACTGGGAGTGCACGTGTGGGCTTTGGTG	2829
Dd	2213	ctgggcaagttcactacagaagtgatgtgtggtcctttggg	2255
<hr/>			
RESULT 14			
ID	AAQ64158		
XX	AAQ64158 standard; DNA; 2128 bp.		
XX	AAQ64158;		
XX	03-FEB-1995 (first entry)		
DE	Partial coding sequence of tyrosine kinase receptor protein.		
KM	Tyrosine kinase; receptor; proto-oncogene; trk; detection;		
KW	diagnosis; antibody; treatment; tumour; antisense; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1..1953	
FT		/tag-a	
FT		/product=	Partial sequence of protein tyrosine kinase.
XX	DE4239817-A.		
XX	01-JUN-1994.		
XX	26-NOV-1992;	92DE-4239817.	
XX	26-NOV-1992;	92DE-4239817.	
PA	(CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.		
PI	Holtfich U, Ruebsamen-waligmann H, Strebhardt K;		
DR	WPI; 1994-184380/23.		
P-	P-PDB; AAR54089.		
PT	New protein tyrosine kinase and related nucleic acid - vectors,		
PT	transformed cells, etc., useful for diagnosis and treatment of		
XX	tumours		
PS	Claim 4; Page 8; 9pp; German.		
CC	The gene is related to the trk proto-oncogene. Antibodies against		
CC	the encoded polypeptide are useful for diagnosis and for the		
CC	treatment of tumours. The antibodies may also be radiolabelled or		
CC	coupled to a cytotoxin for destruction of cancer cells. Antisense		
CC	nucleic acid can be used to inhibit gene expression.		

[illegible]

```

Db      822  tgactacagagaccatccagcgtgatacgaaactccacagaatttgctccaggaggagga 881
QY      1961  TATGAGACCTGAGAGAGCAGCGCCCGCTTCTGCCCCACCTCCCGAAGACGGCTCC 2020
Db      882  ggaagtcagcgtcagcggtgtgtgaagccagtcacgcccagtcgacctgaggggtgc 941
QY      2021  CCATTATGCGGAGCTGACGATTTTACCTGTCAGGGGCTCACCGGGGCAACACCTATGC 2080
Db      942  ccaetacagcagaggtctacatagtgaaactcccaagagatgcagaggaacaacatactc 1001
QY      2081  TGTGCTGCACTGCCCCAGGGGCAATC---GGGATGGGCCCCCAGAGATGATTTCC 2137
Db      1002  agtgcctgcgtacacatgacacgtcctcgaagaagatgtgtgtgtgtgtgtgtgtgtccc 1061
QY      2138  TCGATCTGACCTCCGCTTCAAGAGAGCTTGGCGAGGGCCAGTTGGGAGGTGCACCT 2197
Db      1062  caggaactccctaaccttcaagaagcctggagaaagacagtttggggaggttcactct 1121
QY      2198  GTGAGAGTGAGACAGCCCTCAGATCGTGACAGTGTGATTTCCCTTATATGCGCTAA 2257
Db      1122  ctgtgaagtgagaggaaatggaataatcaagaagatttgcacctagatgtcagtlgc 1181
QY      2258  GGGACACCTTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2317
Db      1182  caacagcctgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1237
QY      2318  CTCTCTCTTCTTCCAGAGATGATTTCTGAAAGAGGTGAGATCATGTGAGGCTCAA 2377
Db      1238  -----ccagaatgatttcttcaagagataagatacatgtctgcgtcaa 1283
QY      2378  GGAOCCCAACATGATTTGGGCTGCTGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2437
Db      1284  ggaacaaacatacatcatatcatatcatatcatatcatatcatatcatatcatatcatat 1343
QY      2438  TACTGATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2497
Db      1344  caetgaatacattgagaaatgagatcgaatcagttcttcccgccagcagggcccttaa 1403
QY      2498  CAAAGCAGCCGAGGGGGGCCCCGTGGGAGCGGAGGCTGCGAGGGGCCACCATCAGCTA 2557
Db      1404  ttcttcctccagcg-----atgtagcaactgtcagttta 1456
QY      2558  CCCAATGCTCTGCAATGTGCGACGCCAGATCGCCTCCGCGATGCGCTATCTGGCCACAT 2617
Db      1437  caccacatctgaattatgttaccacaattgctcctgagatgaagaaacttctctct 1496
QY      2618  CAACTTTGTATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2677
Db      1497  taatttgcacacagagatctgtgcacaaacgaactgtttagtggtaagaactacacaaat 1556
QY      2678  CAAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2737
Db      1557  caagatagctgactttaggaatgagcaggaactgtacagagtggtgacattacacgaatca 1616
QY      2738  GGGCGGGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2797
Db      1617  gggccggagcagtgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1676
QY      2798  CAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2857
Db      1677  caetacagcaagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1736
QY      2858  TAGGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2917
Db      1737  tcaagacagcctcctatccacagctgtcagatgacagagtgatgagatactgtgagagtc 1796
QY      2918  CTTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2977
Db      1797  ctcccgagacaaaggagagaaacttaccctcctcaacagcattgtctcgtactcgt 1856
QY      2978  ATATGAGCTGATCTCGGTGCTGAGAGCCGGAGGTCTGAGCAGAGAGACACCTTTTCCA 3037

```

```

Db      1857  gataagctgatagtcagcgtcgtgagagagagatacagaagaccgtccctcattccaaga 1916
QY      3038  GCTGCATCGGTCC 3052
Db      1917  aatccacctctgtc 1931

RESULT 15
AAC00624
ID  AAC00624 standard; cDNA; 408 BP.
AC  AAC00624;
XX
XX
XX  06-OCT-2000 (first entry)
DT
XX
XX  Human secreted protein 5' EST, SEQ ID NO: 622.
DE
DE
XX  Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW  gene therapy; chromosome mapping; ss.
XX
XX  Homo sapiens.
OS
XX  EP1033401-A2.
PN
XX  06-SEP-2000.
PD
XX  21-FEB-2000; 2000EP-0200610.
PF
XX  26-FEB-1999; 99US-0122487.
PR
XX  (GIST ) GENSET.
PA
XX  Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX  WPI: 2000-500381/45.
DR  P-PSDB: AAG00618.
XX
XX  New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT  obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT  diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX  Claim 1; SEQ ID 622; 71pp + CD-ROM; English.
PS
XX
XX  The present sequence is one of a large number of 5' ESTs derived from
CC  cDNAs encoding secreted proteins. An ORF has been identified within the
CC  sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC  derived from 30 different tissues. EST sequences usually correspond
CC  mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC  often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC  well suited for isolating cDNA sequences derived from the 5' ends of
CC  mRNAs and even in those cases where longer cDNA sequences have been
CC  obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC  mRNAs with intact 5' ends and can therefore be used to obtain full length
CC  cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC  gene therapy and chromosome mapping procedures. They are used to obtain
CC  upstream regulatory sequences and to design expression and secretion
XX  vectors.
XX
XX  Sequence 408 BP; 71 A; 113 C; 126 G; 85 T; 13 other;

Query Match      8.2%; Score 324.4; DB 21; Length 408;
Best Local Similarity 95.3%; Pred. No. 9.1e-62;
Matches 344; Conservative 8; Mismatches 7; Indels 2; Gaps 2;
QY      278  GAGATGCTGCCCCCAGCCCTTAGGCCCCGAGGATCAGAGACTATGGACCAAGGCCCT 337
Db      50  gagatgtgcgtcccccaccccttagggccgagagatcagagcatgagacagagagccct 109
QY      338  GTCATCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 397
Db      110  gtcatcttactgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 169

```

